Keep It Clean

Optimizing EHRs Starts with Ensuring Data Quality

Cleaning up MPI data helps pave the way for successful EHR systems.

by **Beth Haenke Just**, MBA, RHIA, and **Katherine Lusk**, RHIA

If you build an electronic health record (EHR) system, physicians will certainly come. But if you don't build it well—with accurate information and the proper fail-safes to ensure ongoing accountability and effectiveness—they will certainly leave again.

Physician adoption is the key to the success of enterprise EHR systems, and getting doctors to habitually use a system is dependent on one main factor: the integrity of the information in it. That's why a vital step in EHR implementations is cleaning the data in master patient index (MPI) files before linking records in a new EHR.

The Right Technology

Of course, accurate information is vital for everyone accessing an EHR, from the triage nurse in the hospital to the radiologist to the operating room and beyond. It is key to safe, expedited

healthcare. It's not a "wait-and-see" step. It's a "must-do."

In order to have a successful EHR rollout, you need the right combination of people, process, and technology. If you are missing the right people leading the charge, or your people don't have the right skills or they haven't been trained properly, you will have problems. If you don't understand the current process and understand what needs to be modified to get the biggest benefit out of your EHR integration—and put those process changes into place—you will have problems.

And you need the right technology. Not all systems on the market provide a way to do what seems like an obvious step in this healthcare records overhaul—identify errors in the MPI that compromise data integrity and ultimately EHR success.

When it comes to selecting an EHR, one of the things often overlooked is the search function—what method does the system use to locate a patient? Most matching algorithms are good at identifying exact record matches. They fall short when asked to identify records that are "almost the same" or "very similar."

Consider this: humans have a high tolerance for variation. We see things that look alike and are readily able to see the similarities and differences. We can see, for example, that while the names "Damianakis" and "Damanikis" do not match, they look very similar. That ability alerts us to the possibility that the discrepancy may be the result of an error in how the names were recorded.

Computers are a different animal. There is no "almost" or "could be" in computer software. Everything is black and white, bits and bytes. A computer can say with certainty when things are equal and when they are not: "Damianakis" = "Damianakis" and "Damianakis" ≠ "Damanikis."

This is true of the search functions that hospital information systems generally use. These systems do not tolerate a wide range of input errors. If someone spells a name wrong, the record is now wrong. The next time the patient comes to the hospital, even when the registrar does everything right, she won't find the patient because the record on file is inaccurate due to the previous misspelling.

It's in the Search

Basic search algorithms don't do a good job of identifying those records that are close or similar. To detect possible duplicate records requires advanced, mathematics-based search algorithms.

A more sophisticated type of search algorithm uses fuzzy logic and rules-based methodologies. Google provides a common example. Type "fuzy logic" in a Google search, and the system will ask, "Did you mean: *fuzzy* logic?" With this type of low-resolution solution, a patient search function can capture some of the more common errors and thus identify possible existing records—before the problem is further exacerbated with the creation of a second record.

There are even more advanced data-matching algorithms that can be integrated with EHR systems. These algorithms close the gap between computer and human and perform a more thorough search. They incorporate mathematics, statistics, computer science, artificial intelligence, and pattern recognition. Then, modeling the human notion of similar and the human decision process, they return results that are more intuitive to the user and more accurately identify the record the user is requesting.

These systems are able to consider a misspelling, see that there is a nickname, notice a transposition in the Social Security number, or determine that the birth date is off by two years. They can even identify a possible match in a record containing multiple errors. These advanced algorithms are the ones necessary to clean MPIs.

Ten Steps to Record Reconciliation

Ideally an organization performs a comprehensive analysis of its MPI several months prior to launching its EHR. This allows plenty of time to resolve any data integrity issues that may compromise use of the EHR. Cleanup projects typically run three to four months per facility, although the length is dependent upon the volume of duplicates that need to be reconciled.

Careful planning and thorough documentation are critical to a cleanup's success. Documentation should include decisions made on what was and was not cleaned—how many years back, the systems analyzed, and which hard-copy records will be corrected. Record survivorship rules are important to document, as in some cases they can be complex.

It is also important to carefully ascertain the associated risks. An organization may decide not to clean all historical MPI data, for example, yet plan to convert all historical data into the new EHR system. What will be the impact on the system if dirty data are loaded? Another significant risk could be failure to correct paper records if a readily accessible cross reference to the old medical record is not available.

MPI cleanup projects typically follow the process outlined below. The steps apply whether the organization undertakes the cleanup in-house or outsources the work.

1. Finalize the project scope. If work is outsourced, complete the contract.
2. Define the data to be included for analysis; complete the data extract.
3. Complete external data analysis, which identifies all data integrity issues, duplicate records, and overlay records.
4. Kick off the data reconciliation project with presentation of the data analysis results.
5. Conduct on-site project planning to include tasks such as identifying business rules for medical record retention, validity of duplicate decisions, data gathering for procedures to be followed in reconciling duplicates or other data integrity issues, and finalizing procedures and timeline.
6. Complete validity review of all possible duplicates identifying those requiring more research and those that can be successfully confirmed as true duplicates.
7. Complete verifications on all possible duplicates requiring additional research.
8. Complete the merges of confirmed duplicates on all systems or records included in the project scope. This usually includes

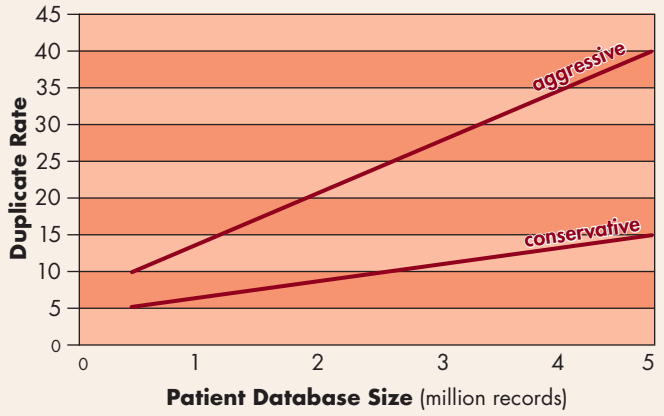
A Visit to the Hospital

Single Data Errors Ripple through an Enterprise's Databases

Let's say a man comes to the emergency room and complains of headaches. He insists he has never been in the hospital before. The staff member inputting his name makes a typo. She doesn't find an existing record, and so she creates a new one. When radiology staff later access the record, they are unaware of the patient's existing files. They do not see that this patient is a repeat visitor to the hospital with the same complaint. A repeat CAT scan is performed—wasting time, money, and exposing the patient to unnecessary radiation.

When staff generate errors and variations in patient data, duplicate records are formed. These mistakes multiply, because the wrong information is accessed repeatedly and new errors are created. Worse yet, the larger a facility's database, the higher the percent of duplicates. This is because a growing database implies more human interaction, which leads to more data errors and variations. Interfaces that automatically link records further aggravate the problem. As the chart below illustrates, a database of fewer than one million records has an estimated duplicate rate of 5 to 10 percent; by the time a database reaches five million records, that rate can reach an estimated 15 to 40 percent.

Duplicate Rate versus Patient Database Size



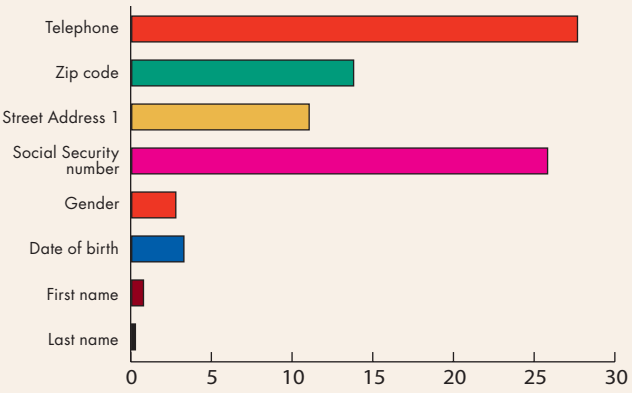
Source: Netrics
As database size increases, the percentage of duplicates increases.

The bottom line is that a core MPI record feeds into many different databases in one healthcare system. A duplicate in the core system propagates throughout the system, such as in radiology, where films or images are stored by medical record number.

The next graphs show how often mistakes happen. These graphs aggregate data across millions of MPI records analyzed at hospitals and clinics across the country. The first graph shows how frequently the data stored in a particular MPI data field are invalid; that is, how often the data field is populated with the incorrect type of data or is left blank.

MPI fields that are populated with a valid value may still contain incorrect data. More than 80 percent of duplicate records contain a discrepancy in one or more of the key patient identifying fields, including name, date of birth, gender, and Social Security number. The last graph shows how often data differ between confirmed duplicate records. (The data shown are compiled across multiple client databases and only include confirmed duplicates.) This, in essence, indicates the cause of the duplicate. Either the data were entered incorrectly on the existing record or the registrar input incorrect data during a subsequent visit, thus generating a duplicate record. And, of course, sometimes, patient data changes, such as last names.

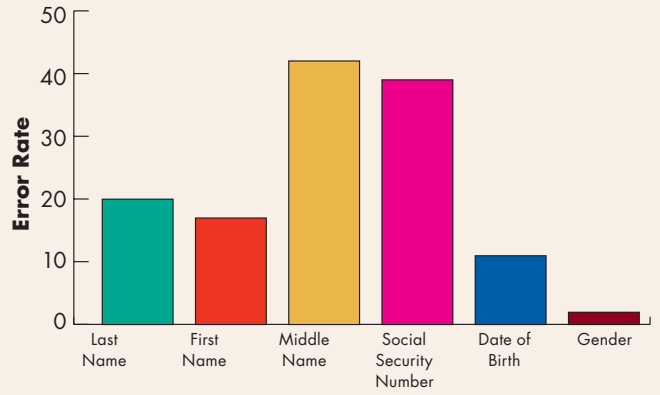
Frequency of Invalid Data Fields



Source: Just Associates, Netrics
Data fields for telephone and Social Security numbers are the most likely to contain null, invalid, or default data.

This is why it is so important to have good process in place and to train staff well to ask the right questions when scheduling or registering a patient. Questions such as "Have you ever been a patient at any facility in our organization before?" or "Have you ever been to our organization under any other name?" are critical to catching name changes. The best software in the world might not catch this.

Fequency of Discrepant Fields in Duplicate Records



Source: Just Associates, Netrics
Middle name and Social Security number are the most likely key fields to contain discrepant data in a duplicate records.

What many healthcare organizations do not realize is that nearly 40 percent of duplicate records contain more than one error or discrepancy. That means that on average at least 3 to 4 percent of all records in an MPI database have multiple errors. That may not seem like a lot in the scheme of things, but as one pediatric surgeon said, "One is too many if it's your child or your patient." ❖

the main hospital or clinic registration system and medical records corrections and frequently includes radiology, laboratory, and other clinical department systems or records.

9. Summarize project findings, including causes of data integrity issues. Provision of maintenance strategies and complete documentation of knowledge gained throughout the project helps ensure the facility can keep its database free of duplicates and other data integrity issues.
10. Complete a project evaluation to ensure understanding of project successes and opportunities for improvement.

It Works!

It worked at Children's Medical Center Dallas. There, the organization's executives and physician leadership recognized their EHR investment was useless if the quality of patient-identifying data was not excellent.

Children's had loaded data from legacy systems into its new hospital information system, had interfaces creating new records in this system from its laboratory system (where it registered specimens), and was running a new registration system that functioned quite differently from its old system. The hospital had been imaging its medical records for a few years and was about to launch a major electronic clinical documentation initiative.

The HIM and IS directors agreed that they needed to clean up the MPI data and ensure that old issues did not surface again. They set out to assess their current registration and registration training processes, evaluate system configurations, analyze the MPI, and then clean up the data.

They knew this was a critical step in ensuring the success of the EHR. A hospital survey found that many doctors could not locate patient information when needed:

- →45 percent found duplicate medical record numbers
- →25 percent felt the duplicate rate affected quality of care
- →30 percent believed they had re-ordered tests due to duplicate medical records

The fiscal impact of these survey results, not to mention patient care risks and physician frustration, was enormous. A cost study of 1,000 confirmed duplicate records evaluated each duplicate to identify repeat tests, delays in treatment, or other quality of care issues. The study demonstrated that on average, a duplicate medical record cost an organization more than $96. This excludes the cost to correct the duplicate but includes costs such as increased registration time.

More than 4 percent of the confirmed duplicates resulted in some type of clinical impact. The most common issue identified was delay in initiating treatment in the ER. Other common quality issues identified included:

- →Duplicate test ordered due to lack of access to previous test results. Most commonly these were radiology exams where the previous film was not available for comparison.
- →Delay in surgery due to a lack of a history and physical at time patient presents for surgery.
- →Pre-op visit on one medical record number and then surgery under the duplicate medical record number.
- →Transport records under one medical record number and hospital stay under second medical record number.

Morgan Stewart
HEALTH INFORMATION SERVICES

Quality HIM Services you can depend on!

Commitment to excellence and first-class customer service are at the heart of Morgan Stewart. Contact us today to see how we can become an integral part of your HIM team!

- Coding (Local, Travel & Remote)
- Inpatient & Outpatient Coding Quality Reviews
- Revenue Cycle Enhancement
- Interim Management
- Clinical Documentation Improvement Programs
- Coding Education Programs
- JCAHO Prep Programs
- Opportunities Available Nationwide

www.msqim.com
1-877-877-5802
or charlie@msqim.com

Quality Information Management
Quality in Motion!

AssistMed®
Unleashing the Power
Of Healthcare Information™

Why learn an EHR?
An EHR can learn from *you*!

Intuitive
Dictation-centric
Speech Recognition
NLP Knowledge Extraction
Customized Workflow Scenarios

Online, Subscription-based Solutions
No Capital Investment
No Maintenance Costs
No Obsolescence

www.AssistMed.com

PROFESSIONAL DYNAMIC NETWORK, INC.

PDN

Personnel Staffing
Healthcare and Government
Temporary & Permanent

Certified MBE/WBE

Specializing in
Health Information Management
(All Medical Record Services)

PDN serves hospitals, clinics, physician groups, government agencies and others

Clerical
Technical
Administrative
Managerial

Serves:
Chicago, Suburbs, and NW Indiana

Call 708-747-4361
www.pdnseek.com • info@pdnseek.com

Children's hospitals face unique challenges in patient identification because their patients do not have the standard forms of identification. This creates an even greater focus on staff training, scheduling, and process design. Children's took a comprehensive approach to rectifying its data integrity issues. Some of the key actions in its successful approach included:

- Forming a multidisciplinary committee to oversee the evaluation and corrective processes
- Ensuring high visibility of the project to executive and physician leadership
- Bringing in a consultant with expertise in MPI data integrity and in the hospital's new information system
- Thoroughly analyzing all data integrity issues, quantifying the causes of each and generating action plans to address every cause
- Analyzing user creation rate patterns to identify those staff that required more training
- Evaluating interface configurations for feeds into and out of the system
- Completing a comprehensive process review of all registration and scheduling areas, including review of policies, procedures, and training manuals
- Evaluating current process for reconciling duplicate and overlaid medical records

The result was a comprehensive plan that enhanced the registration and scheduling training program, reconfigured the registration system, increased data record match requirements to prevent overlays, cleaned all duplicates, and implemented an enterprise MPI with advanced search algorithms. Children's then developed a process to continuously evaluate data integrity, and it dedicated appropriately trained personnel to monitor the system.

The effort showed meaningful results. Before the cleanup, the volume of errors was so high that the hospital had set up a dedicated number to receive reports of duplicates in the system. After the cleanup, the phone stopped ringing, e-mail stopped arriving, and the staff stopped complaining. The rate of duplicate medical record creation dropped to 0.2 percent. Physician complaints stopped.

Keep on Keepin' On

Once duplicates have been removed, information is clean and accurate, and the system is operating effectively, the job is still far from over. Staff must remain vigilant to keep problems at bay.

If errors pop up, look for patterns. Find out if the problem is being created in a specific situation or in a specific department. If you find a problem, solve it. Offer more training. Put policies in place and monitor their use. Don't allow duplicates to get away from you. Stay on top of them to optimize your EHR investment and you'll be hitting satisfaction home runs with physicians—and patients—for a long time to come. ❖

Beth Haenke Just (bjust@justassociates.com) is founder and CEO of Just Associates. **Katherine Lusk** is director of health information management at Children's Medical Center Dallas and an adjunct professor in the health information technology program at DeVry University in Irving, Texas.